A Work Project, presented as part of the requirements for the Award of a Master Degree in Finance from the NOVA – School of Business and Economics.

The impact of big data and data analytics in external auditing procedures

Joana Pedrosa Santos (student number: 24074)

A Project carried out on the Master in Finance Program, under the supervision of:

Associate Professor: Maria João Major

4th of January, 2019

Abstract

Companies that can effectively utilize big data, on average, have a 5 to 6 per cent higher gain in productivity (Brynjolfsson et al., 2011). This fact is not going unnoticed by the auditor's community that has been implementing data analytic tools in their projects. The aim of this working project, written in the form of a direct research on EY assurance department, is to understand this impact on auditing procedures.

Facilitators and obstacles are the factors that determine how big data and data analytics usage will evolve in audit practice, with leadership commitment identified as the most significant facilitator and aversion to change as potentially the greatest obstacle.

Keywords: Big data, Data analytics, Impact, Auditing procedures, Leadership

I. Introduction

The main research question this Work Project (WP) seeks to answer is what is the impact of big data and data analytics in audit procedures. More specifically, by conducting interviews to auditor professionals and a four-month effective work on the EY assurance department, a Big 4 company, the study purposes to disclose how auditors in these companies perceive the utility of big data and data analytics to support auditing procedures.

The topic was triggered by the crescent awareness of accounts, audit firms, standard setters, and academics (Rezaee et al., 2017). For instance, American Institute of Certified Public Accountants' Assurance Services Committee (ASEC)[1] has met Auditing Standards Board (ASB) for several times to develop a new AICPA Audit Data Analytics Guide. Also, all Big 4 have made announced efforts in the data analytics field (Deloitte, 2013; PwC, 2018; KPMG, 2018; EY, 2015). In addition, International Auditing and Assurance Standards Board (IAASB) is developing the Data Analytics Working Group (DAWG)[2] with the aim to inform stakeholders about the IAASB's ongoing work to explore effective and appropriate use of technology, with a focus on data analytics, in the audit of financial statements. Likewise, the PCAOB has currently approached the Big 4 to discuss the data analytics.

This dissertation is structured as follows: Section II reviews the literature on big data and data analysis in audit procedures taken into consideration by the researcher; Section III describes the methodology followed; Section IV presents the empirical study and discusses the main findings and finally in Section VI the conclusions and main contributions of this WP are presented.

[1] The ASEC's Emerging Assurance Technologies Task Force has worked on three projects related to audit analytics: 1) creating audit data standards, 2) developing a framework that maps innovative data analytic procedures to the more traditional audit objectives and 3) developing a dataset with essays that discuss continuous assurance, continuous control monitoring, and advanced analytics techniques.

[2] Available at: http://www.iaasb.org/projects/data-analytics

II. Literature Review

The auditing profession is being challenged by advances in information technology allied with real-time approaches to manager businesses. Businesses are using big data and business analytics to help them face ever-increasing business complexity (Rezaee et al., 2017). In fact, Brynjolfsson et al. (2011) argue that companies that can effectively utilize big data, on average, have a 5 to 6 per cent higher gain in productivity. At the same time that there is a clear recognition of the importance of data analytics (Vasarhelyi et al., 2015), the audit community also faces challenges and questions about the achievability of a full integration on the auditing procedures.

Big data is typically described as 4 Vs: volume, velocity, variety and veracity (Buhl et al., 2013). Moreover, it is often referred to as enormous datasets consisting of unstructured, semi-structured, and structured data collected from several sources that can be processed electronically in real time (Cao et al., 2015). Data analytics are tools that allow to analyze and transform this vast amount of data into structured information in a timely manner for decision making (Tang et. al., 2017). Big data and data science principles enable auditors to capture data by gaining access to both internal and external parts of to their client's company (Rezaee et al., 2017).

Vasarhelyi et al. (2015) argue that big data has changed the manner that auditors acknowledge information in the same way that it has already influenced business practitioners. Schneider et al. (2015) realize that big data together with data analytics, impact on auditing by reducing the cost of gathering, processing, and storing information, what, in last instance, means to enhance audit firms profitability (Littley, 2012). Furthermore, these tools have the potential to lead to better use of the information to interpret, estimate, and assure their processes (Dai et al., 2013), predict future tax liabilities (Schneider et al., 2015), detect fraud (Cao et al., 2015), corroborate the going concern prepositions (Littley, 2012) and identify risky transactions (Alles, 2014). In addition, big data will be fundamental to enlarge the sufficiency, reliability,

and relevance of audit evidence, which further improves audit quality (Alles, 2016). Ultimately, Alles (2015) states that auditors may promptly embrace big data as a way of increasing the effectiveness and credibility of their work. After receiving harsh criticism under the Enron and WorldCom' scandals and the financial crisis, auditors feel some pressure to adopt big data and data analytics as a way to ensure additional credibility about their work (Alles, 2015) (see appendix 1 for a summary of the main advantages of using these tools as part of audits).

It is important to answer the question of why auditors would embrace big data when simpler methods have not been embraced. For instance, auditors fail to make full use of the non-big data that they already have access to. In fact, many auditors continue to rely on manual confirmations, using samples when full population data are available, and basing analytic procedures on aggregates rather than the analysis of disaggregated transactional data (Alles, 2015). Alles et al. (2002) made the argument that this is a demand rather than a supply-side argument making auditors turn to big data. In other words, the imperative to keep up with their clients' reliance on big data is the exogenous driver that will force auditors to also adopt big data. In the past, the widespread adoption of Enterprise Resource Planning systems (ERPs) was one key factor that enabled and ultimately made inevitable IT-based auditing. The need to keep pace relative to their clients' overall business environment force auditors to adopt the same tools as they do in order to identify risks. For example, Netflix has been using big data for colour and title analysis in its TV show covers. This data has allowed Netflix to conscientiously understand its customers' preferences[3]. Hence, activities that contradict the established customer preference can be recognized by the auditor as areas of risk. In short, there is a strong belief that the key factor to induce auditors to adopt big data will be the exogenous market forces. Moreover, it will induce the adoption sooner and more fully than it otherwise would be. Besides, the profession's reaction to big data will strengthen auditors to maintain their

[3] Simon P. 2014. "Big Data Lessons from Netflix". March. Wired. Available at: http://bit.ly/2pJLw9W

credibility with their clients who see big data as a strategic driver of their businesses (Alles, 2015).

Auditing is highly constrained by standards and, in the case of external auditing, focused on the specific task of attesting to management assertions. Historical and legal structure constraints have to be taken into consideration when discerning this topic (Alles, 2015). It was expected that the adoption of big data by enabling the induction of nontraditional sources of data into the auditing sphere would result in sharped auditing standards, consistent with the new domains (Vasarhelyi et al. 2015). Nonetheless, regulations, such as the guidance for sampling, have remained unchanged even though many audit clients automate the collection and analysis of 100 per cent of their transactions (Schneider et al., 2015). Still, it is perhaps unwise to assume standards as being a limitation to the use of big data since auditors are currently not employing the full spectrum that those standards allow. Alles (2015) made the argument that instead of creating new or modified standards to boost the use of big data and analytics methods, auditors would quickly accept these tools if they are seen as a means of more effectively and/or efficiently satisfying existing standard.

Returning to the principles topic, the primary objective of an external auditing engagement is to give an opinion regarding the client's financial statements and internal controls in the auditor's report (Appelbaum, 2016). In order to reach that output, auditors have to gather sufficient and reasonable evidential assurance that those financial statements are relatively free of material misstatements. This implies that auditors should plan and then complete procedures to collect enough verifiable evidence (ibid).

Looking for international auditing standards, ISA 200 states that an auditor must gather "appropriate audit evidence (...) to reduce audit risk to an acceptably low level, and thereby enable the auditor to draw reasonable conclusions on which to base the auditor's opinion." Moreover, ISA 200 paragraph A28 specify that such audit evidence can be from "sources inside

and outside the entity" and that the purpose is for the auditor to find data that "supports and corroborates management's assertions". However, Sarbanes-Oxley Act (SOX) demands that auditors authenticate the veracity of the evidence that supports their audit opinion (Appelbaum, 2016). This methodology has been influencing the way audit procedures are conducted by impacting the profession's analytical procedure choices. Substantive tests, which depend more heavily on detailed audit examination had more power over regression and other "softer" analytical techniques (Glover et al., 2015). Given the response to these concerns, ISA 200 paragraph A29 presents a doctrine that could be the key to use big data to support audit' work: "The sufficiency and appropriateness of audit evidence are interrelated. Sufficiency is the measure of the quantity of audit evidence. The quantity of audit evidence needed is affected by the auditor's assessment of the risks of misstatement (the higher the assessed risks, the more audit evidence is likely to be required) and also by the quality of such audit evidence (the higher the quality, the less may be required) sets in order to achieve sufficiency". Following a similar principle, ISA 330 requests the analysis of the source and reliability of the available information and the persuasiveness of audit evidence. Overall, these standards capture the big picture of what constitutes audit evidence and they are not written in a way that constrains the use of any source of data (Rezaee et al., 2017). Nevertheless, auditors are lagging regarding the use of nontraditional data to support their work. At the same time, the lack of guidance of the international auditing standards could be an obstacle to the evolution of big data in auditing in the future (Alles, 2015).

While standards do not seem to drive barriers to the adoption of big data (for now), the use of non-traditional data brings other issues. In this sense, evidence collection and analytical procedures should be considered. Since, it is imperative for auditors to test for basic assertions to make sure that the purpose of the audit engagement is achieved regardless of the different nature of evidence and the way this evidence is gathered and analyzed (Appelbaum, 2016).

When evidence is based on big data analyses, Alles and Gray (2014) advert for the difficulties that auditors may have with the lower standards of data integrity. Under these circumstances, it would be advisable for auditors to at least validate the data treatment and perhaps to replicate that analysis independently (Alles, 2015). About this subject, Cukier and Mayer-Schoenberger (2013, p. 28) state: "Using great volumes of information in this way requires (...) [Big Data users] to shed our preference for highly curated and pristine data and instead accept messiness: in an increasing number of situations, a bit of inaccuracy can be tolerated, because the benefits of using vastly more data of variable quality outweigh the costs of using smaller amounts of very exact data". This judgmental exercise comes with the introduction of news, audio and video streams, cell phone recordings, social media comments as data to obtain new forms of evidence, confirm the existence of events and validate reporting elements (Moffitt et al., 2013). For these reasons, it becomes crucial to test the quality and verifiability of these external non-quantitative sources for the process of risk assessment evaluation (Appelbaum, 2016). For instance, clients can use external social media sources in order to do financial valuations of some assets (ibid). Unfortunately, the reliability of tweets and other external social media is difficult to confirm (ibid). Adding to that, PCAOB (2010) states that if the collected evidence has unreliable origins and attributions then it is not possible to rectify that dataset by gathering more of such data.

The introduction of big data is a turning point regarding concerns that the different sources of information require (ibid). It is totally different talking about electronic accounting and audit information and manual and paper-based sources. Since many of the positive attributes that paper-based evidence present are now challenged by electronic evidence. While paper documentation is usually difficult to change, the biggest issue about digital data is how easily it can be altered and how undetectable these modifications can be. This raises the problem of lacking security control over the data provided (ibid). Besides, paper-based evidence brings

the additional security over the reliability of third-party external sources (PCAOB, 2010), whereas with third-party data from several sources, it is challenging to establish origins and veracity (Appelbaum, 2016). Consequently, the types of tests that should be performed by auditors to examine basic assertions may change or at least the control over the data used.

Even though big data can carry some risk, one has to balance between detailed evidence collection and analytical procedures (Yoon et al., 2015). Which could be not so black and white. The standards clearly require the performance of tests of details of transactions; however, the requirements for the use of analytical review procedures is somewhat undefined, except that it should be applied to financial data (Rezaee et al., 2017). The adoption of one method or another is at the end of the auditors' professional judgment. Nevertheless, there is a trade-off between these two kinds of approaches. Detailed evidence collection can assure more reliable data whereas it is also costly. In fact, analytical procedures are broadly viewed as being less costly, yet less reliable by regulators (Rezaee et al., 2017).

According to the auditing method, there are four phases of an audit engagement. The introduction of big data and data analytics on the audit process triggers the question of to what extent these methods can be included and in which phase. In the planning and design audit phase, there are some audit activities that are likely to benefit from big data analytics. Firstly, it can help to identify and evaluate risks associated with taking or continuing an audit engagement (Rezaee et al., 2017), when facing risks of bankruptcy or high-level management fraud. Secondly, it is also useful to identify and assess the risks of material misstatement for fraud (ISA 240) or even for understanding the company, its environment and design further auditing procedures shaped to the identified risks (ISA 315). Given the wide variety of sources, the increased use of big data holds much appeal as a way to improve the effectiveness of preliminary analytical review procedures (Rose et al., 2017).

The second phase of the audit engagement includes performing preliminary analytical procedures, as well as evaluating the design and implementation of internal controls and testing their operating effectiveness. In this sense, big data is useful to gather and inspect data for evidencing that control is operating as it is supposed to. Additionally, data analytic tools enable auditors to search for patterns in big data that would likely be undetectable in typical audit samples (Alles and Gray 2014; Alles 2015). In some circumstances, data analytics can also rehash the control activity itself.

Analytical procedures are used in the substantive testing phase to obtain evidence about assertions related to critical accounts or business cycles identified based on auditor's assessment of the risk of material misstatement (ISA 520) (Appelbaum et al., 2017; Cao et al., 2015). Data analytics can be used to perform a recalculation of an entire population as opposed to only a sample of items. For instance, data analytics can be used to perform reconciliations by comparing information from several sources. Yoon et al. (2015) reinforce the introduction of big data in this phase arguing that analytical procedures may be more effective than tests of details in some circumstances.

For the review phase of the audit engagement, analytical procedures are required to evaluate the auditor's conclusions regarding significant accounts and to support the audit opinion (ISA 520). Nevertheless, it is the phase where big data will have less impact since the conclusions of an audit report are mainly based on professional judgement.

The implementation of big data and data analytics in audit is not easy. It can raise several issues. Firstly, auditors face a paradigm shift between using only some clean data to use "all" the data in large relatively messy datasets. This implies a change in the focus of cause to correlation (Cao et al., 2015). Furthermore, the degree to which these tools can be applied depends on the stage of the audit. For example, planning and risk assessment phases easily tolerate the use of unstructured data, on the contrary, substantive procedures are more sensitive

to the organization of data. New guidance and education on audit field linked to automated systems is required since the integration of big data demands a greater skill set (Cao et al., 2015). Competencies that audit professionals do not have it yet.

Secondly, big data represents a challenge in practical terms since the huge volume of data brings computational issues. Right now, the analytics tools used by auditors do not have the capability to take advantage of big data. In this case, the solution may involve selecting subsets of data to be managed by more complex analytics tools and, simultaneously, assure the availability and relevance of the information collected and merged (Rezaee et al., 2017). Nonetheless, data aggregation is complex mostly due to data incompatibility since it comes from several sources. Settles (2009) made an argument supporting this point: he states that the selection of subsets of data results in more accurate analytical models.

Thirdly, privacy is a potential concern when big data is used, namely, regarding data storage (Cao et al., 2015). Although, this risk is not specific for auditing it represents a significant challenge both for clients and regulators. For instance, analytics tools can require the use of clients' nonpublic information beyond the data usually available to auditors or even of information from previous audits to improve its performance, perhaps from other clients (Cukier, 2013). Following this same logic, external auditors' independence might be questioned when they have too much information about their clients.

Lastly, big data and data analytics application comes with the risk of not detecting a fraud or error. This is not necessarily a new problem as until now auditors have traditionally based their work on samples, and it is accepted that there is a statistical risk for fraud or error not to be detected. However, when auditors apply big data, it is assumed that the hypothesis for this situation to occur is residual or even null (Cao et al., 2015).

In another perspective, the emergence of big data also carries many new opportunities. For instance, there is a big potential to apply time series in auditing. Time series models can be

used to convert unstructured data sources into a sequence of points, typically over a period of equal time intervals in order to discover a pattern in the data (Rezaee et al., 2017). One of the primary applications of time series is in forecasting and budgeting in areas such as sales revenue, operating income, net income, and stock prices. It represents an improvement compared with the current practice that ignores factors as seasonal, cyclical, and irregularity components. Through the use of sophisticated time series models, some conventional auditing procedures are outdated. Auditors used to develop hypotheses based on expectations and then to collect data to support or reject the hypotheses. As opposed to traditional statistical analysis that uses data to test hypotheses, time series can be embedded into the system and detect unexpected material deviations and red flags whenever there is divergence from what is determined by the model to be the norm or acceptable (Rezaee et al., 2017). Thus, big data and time series allow auditors to perform this job with considerably less cost and in a shorter time period compared to traditional models.

In his article: "How Big Data and Analytics are Transforming the Audit" [4], Ramlukan (2015) claims that, perhaps, the most significant contribution of emergence big data is its ability to provide population-based audit evidence. For instance, the full access to the accounts receivable of the client´s records would permit to mitigate the bias from sampling. Moreover, it will allow further analyses by slicing the accounts receivable in different variables (e.g., transaction amounts, time, location) and make comparisons across the groups to find patterns and obtain more meaningful insight.

Big data also can be applied to strengthen the degree of prediction accuracy (Tang et al., 2017). This feature is useful not only to build connections between financial accounts based on real-time information, but also to predict the relationship between a company's financials

[4] Available at: http://bit.ly/2pJJ8ju

and industry averages. Crossing financial and nonfinancial information could be a great improvement in fraud detection. This is especially relevant to monitor management and boards of directors. With text-mining techniques, it is possible to use emails, phone calls and even meetings to collect information and identify hidden patterns or links within financial data (Mayew et al., 2012).

Finally, big data can also be used in the auditing of external business relationships (EBR). Big data enables an auditor to gather information on a client's EBRs, particularly in risky areas not captured by financial data. For instance, reputation damage detected on an online review or new reports of a client' supplier can harm the business and consequently affect revenues.

Moreover, audit firms can potentially use big data for internal purposes. Considering that almost all auditing evidence are in electronic form, there is an opportunity for the audit firm to analyze audit engagements across an entire portfolio in search of anomalies and potential quality issues (Cao et al., 2015).

Overall, the emergence of these analyses brings four main consequences: (1) descriptive statistics on the entire population; (2) audit evidence on a larger and more complete scale; (3) connections between financial statements and actual business operations; and (4) identification of potential red flags (Tagan et al., 2017).

III. Methodology

Taking as reference the methodology described by Ryan et al. (2002), after reviewing literature and identifying its gaps, research questions should be formulated. The research question that this WP aims to respond is *"How do auditors in Big 4 companies perceive the utility of Big Data and Data Analytics to support auditing procedures?".*

The research method adopted was a case study, based on a qualitative approach. Case studies are recommended under three specific conditions: (i) when the object of study is a

contemporary phenomenon; (ii) when the researcher has no control over the phenomenon object of the investigation; and (iii) when the objective of the research is to get a deep understanding of the phenomenon within its context. (Yin, 2014).

This dissertation was developed in an environment that satisfies all the three conditions. In order to be able to cross data from multiple sources, evidence for the investigation was gathered in two different moments. These sources include semi-structured interviews, direct observation and participative observation through a four-month effective work as auditor in the EY assurance department. Interviews enabled the researcher to focus directly on the WP topic and understand in depth the internal perspective and procedures of the company (Yin, 2014).

To get a better perspective of the topic studied, the semi-structured interviews were conducted to people from different ranks within the organization. This allowed to have a heterogeneous group of interviewees. Furthermore, the interviews were conducted during the four-month of the researcher in the organization' framework (between September 2018 and December 2018). The time duration for each interview was 40 minutes. A depiction of the interviews carried out is presented in appendix 2 as well as the corresponding guide (see appendix 3). The interviews were conducted in Portuguese and subsequently fully transcribed to English (see appendix 4).

In this study, the author adopted the same EY organization hierarchy in order to define the interviewees position in the company. The ranks by descending hierarchy order are the following: partner, associate partner, executive director, manager, senior and staff. The author chose employees of different ranks with the objective of reaching a more heterogenous population with a wide span of perspectives.

Interviews started with specific questions on how interviewees perceived the changes on the audit procedures with the implementation of big data and data analytics tools. More general and open questions about the advantages and disadvantages, obstacles and risks

associated with those changes were placed in the second part of the interview. Finally, the interviewees were asked about the impact of audit standards in the implementation of big data and how they perceive the auditing' future.

Direct and participative observations were also employed as they enabled the researcher: to cover events in real time and on the job; to understand the context of the event; and to obtain insights on interpersonal behaviour and motives to adopt and implement big data and data analytics (Yin, 2014). Given the fact that the researcher performed a very active role in the company as an auditor, having the possibility to deal directly with the data analytics tools used and to interact with their coworkers, the author was able to get inside knowledge about the efforts and progresses made on the field. Thus, the author's role in the investigation can be described as an 'actor researcher' (Ryan et al., 2002).

The main findings and conclusions on this WP were produced taking into account the patterns verified through the several sources of evidence (Ryan et al., 2002; Yin, 2014). Patterns were better understood following the preparation of a diagram based on the answers obtained in the interviews (Ryan et al., 2002) (see appendix 5).

IV. Empirical Study

The first point to emphasize, which is common in all opinions of the interviewees, is that the adoption of big data and data analytics in the audit procedures is at the beginning.

Nevertheless, interviewees also observe that audit is changing and is moving fast; accordingly, a growing body of audits are expanding the usage of these tools as part of an audit mainly driven by internal politics of the company instead of external pressure as literature suggested. A partner interviewed (Interviewee 4) said:

> *"I think it was due to the position that EY adopted from wanting to stand out and from wanting to be the first. When we did the Vision 2020, the EY wanted, wants and will always want to be the number one in the world and therefore had to*

> *stand out. (...) So, this approach was in that sense, the will to be the best. The driver, was mainly internal and in the direction of increasing quality"*

This is indeed one of the drivers pointed out by EY auditors: the commitment of the leadership with the implementation of the new technologies. This is not a commitment in vain, steps have been taken to make the implementation of the big data and data analytics real and global to all projects. As it can be concluded from the statement of the manager interviewed (Interviewee 2):

> *"In the evaluation we all have a part in our rate in the last 3 or 4 years connected to what is our performance in the digital viewpoint. It means that if I use the tools in 60 or 70% of my clients, I'm fine. (...) Is the way of leadership to pass the burden of implementing some tools for the ranks of the organization."*

Before assessing the impacts of big data and data analytics, the researcher has identified a gap between the theoretical definition of big data and big data as auditors perceive it. In this matter, it is important to differentiate between more of the same type of data that auditors are already using (traditional accounting data), and in which they usually rely on, or more data of a different source (Alles, 2016). In the first scenario, an auditor could just call on continuous auditing by increasing data volume based on time-frequency instead of expanding the scope of data. On the other hand, big data implies covering not only financial information and structure data from inside the organization, but also non-financial and unstructured data from exogenous sources. Connolly´s (2012) framework (see appendix 6) is useful to approach this topic. It shows in perspective, the data currently used by auditors, in a small box in the lower left corner, in comparison with all the potential information that the big data offers to expand that input. The Interviewee 3 states the following:

> *"When you talk about big data in financial terms, that is, in audit context, it is about to manage all the journal entries (JE's). Collecting information outside the company is almost impossible."*

The main impact of the adoption of big data and data analytics techniques appointed by the auditors interviewed is that with those tools auditing procedures are no longer dependent on sample testing instead they rely on 100% sample analysis which can improve the security of the analysis. Interviewee 4, says the following:

> *"In terms of positive impacts what I think that we have effectively been able to visualize a lot more and perceive the business in a much better way than we realized. That is, being able to visualize how a given transaction starts, processes and ends allows us to have a much more comprehensive view not only at the audit level, but at the level of producing reports for stakeholders. This is undoubtedly one of the great advantages."*

Alles's scheme (see appendix 7) addresses this point in a more visual way. It illustrates how big data and data analytics concepts can be related in the audit domain. While, for many years auditing firms have been comfortably operating in 'Cell A' using traditional data analytic tools, as Excel, to analyze samples of financial data. Now, they have started moving into 'Cell B' by moving away from sampling and starting to adopt data visualization tools, such as 'Tableau'[5].

The point mentioned above necessarily implies a paradigm shift in the way audit processes are conducted. From a sample-driven approach (see appendix 8) to a data-driven environment (see appendix 9) where the examination process would be analytically reviewed, audited automatically, and exceptions or outliers would be subsequently examined in detail. Following this rationale, auditors may shift from an audit engagement typically progresses as shown in appendix 8 to eventually innovate to a more audit by exception (ABE) approach (Vasarhelyi et al., 1991). In ABE an audit process is activated by alarms triggered in data streams. Indeed, this logic is pointed out by interviewee 1:

> *"Theoretically, the big data has the ability to create variable correlations. (…) Putting this in auditory for dummies, (…) If I am analyzing the totality of my*

[5] Tableau is a software focused on business intelligence that allows the interactive visualization of data.

> *sales, I can say that in principle, everything that is sales in return for customers follows a certain pattern and sales in return for cash follow a less common pattern. So, if you have a high sales value against cash I have here a change to the standard that I should analyze. Sales in return for any other heading are clearly exceptions and so I will need to get explanations for these exceptions."*

Additionally, auditors also found in big data and data analytics methods an opportunity to improve quality and efficiency. As well as a way of lengthening their scope of work. By using non-financial data, auditors can perform another kind of analyses that they are not able before. This opportunity raised since auditors could focus on drawing a better conclusion and to add value to their client with better recommendations. Interviewee 4 mentions:

> *"The work itself increased in quality and it was undoubtedly much more secure. We now realize that we have to go looking for other things, not only whether or not there are errors in the accounts is whether someone made a particular transaction. (…) From the moment we stop working with statistics and work as a whole, our audit becomes much more reliable. We can draw much more comprehensive conclusions."*

In this perspective, the associate partner interviewed draws attention to the client's side who starts to see the work of the auditor with more credibility and at the same time understands better the requests for information of the company. However, on the other hand, the client may also want to negotiate the auditing fees claiming that with these tools then an auditor has to be more efficient. As Interviewee 1 contends:

> *"But in the future, then, it is also possible that clients want to negotiate fees in order to lower the cost of the service."*

Following the structure of the interview, some obstacles and risks associated with the adoption and implementation of big data and data analytics as part of the audit engagement will be presented. The technology adoption curve developed by Moore (2002) can be applied to give some context to this matter (see appendix 10). The charms on this figure represent the barriers to the adoption of a new technology between the different types of technology adopters of the

organization. Specifically in the auditing context the first chasm, between the innovators and the technology adopters, can be perceived as the gap between senior-level partners, who speak at conferences, and the frontline auditors. One of the triggering factors is related with the investment needed as interviewee 4 commented:

> "*In fact, we have invested many billions in this kind of tools, in this kind of way of auditing to stand out. Here, we also gain from being global. (...) This helps us have much more powerful tools. It would be impossible for us in Portugal to do all this alone."*

As well as the human resources required to develop and implement new tools. The manager interviewed (interviewee 2) said:

> *"It also implies the greater involvement of people with more experience with the ability to make decisions and complete the work successfully. And the reality is that there is also the lack of availability by everyone to be able to do this kind of work."*

Another factor is related to the need to create user-friendly tools since many of the practitioners do not have enough technical skills to handle more complex software. This factor turns out to be entirely related to the previous one since increased technical skills on the part of the audit team presuppose both short and long-term cost consequences. As the associated partner interviewed (interviewee 1) mentioned:

> *"I have to create tools that are intuitive in processing analysis because, for example, I speak for myself, I know how to analyze the data, but I do not know how to program to process the data and this is also one of the big data challenges."*

Nevertheless, the most challenging chasm is between early adopters and early majority. While the first ones embrace the new technology as an opportunity and seek to potential applications the second group is loath to change and is only able to adopt the new technology when they are convinced of its benefits. Auditors point out the limited time to perform an audit

engagement as one of the reasons to avoid implementing these tools. This explains why they opt to follow previous years models to do certain analyses. Interviewee 2 claim:

> *"A major obstacle to change is people. People are very averse to change. They want to continue doing what they have always done. This is always the biggest obstacle, it is the mentality."*

Making a parallel with the Alles' matrix (see appendix 7), auditors are facing major challenges migrating to 'Cell C' and truly incorporate big data and sophisticated analytics in 'Cell D'. It also can be seen in appendix 1, it means that the auditor will be expanding outside of the small box in the corner into a vast population of nonfinancial data. The access to a large amount of sensitive client data was one of the hurdles identified on this stage since auditors gain extra responsibility to ensure that there are no breaches of confidentiality and privacy. This matter was raised by interviewee 5:

> *"Some problems also arise due to privacy restrictions imposed by the client."*

Other of the risks associated with the mass use of data is, as reported in literature, the quality and veracity of this data. Making the extraction and testing of the data a relevant setup cost for auditors. This is also one of the risks mentioned by the interviewee 1:

> *"The auditor must take the most care in analyzing all the information provided by the client. (...) I have to make sure my starting point is correct. (...) When we are talking about using big data and we use tools that systematize this question, (...), due to its massification this first test done to the IPE[6]s is not made with human or artificial intelligence, reason why, if I have had some flaws in IPE a priori all reviews will be tampered with. This is one of the great risks that we have in this analysis of mass information, (...), we have to audit the initial information that gets into the black box and then is processed."*

Additionally, adopting such disruptive technologies imply a paradigm shift as to how financial statement audits are conducted and a significantly higher reliance on non-financial

[6] IPE means information provided by the entity.

data. Richins et al. (2017) developed a scheme (see appendix 11) that illustrates the contrast between a problem-driven and an exploratory analysis. The dashed arrows that diverge from the two data types to the functional nodes outline the concept of an exploratory analyze. Unlike problem-driven analysis, this kind of approach does not allow strong *a priori* predictions instead practitioners have to discover knowledge encoded in the data. The Interview 3 stated the following:

> *"Before you had to look it up first, but now you don't need to do it, just input the information and the system will tell you where the amounts are going, and you just need to check the counterparts. You don't need to go from JE to JE or looking for invoice."*

However, the feature that has most concerned auditors is that not all of the arrows eventually make it to one of the accounting nodes. In other words, it means that not everything the data has to say will be useful or relevant for decision makers. Thus, as written by Keltanen (2013) the pure extraction of data is prone to create false correlations. Interviewee 1 also states this risk:

> *"I can take the risk of obtaining only a summarization of information exclusively and this is not audit."*

Additionally, and contrary to what Cukier and Mayer-Schoenberger (2013) defended about large data scale might tolerate some messy data. Auditors may have some difficulty to reconcile with the focus in auditing on data integrity. Doing so also goes against the desire of auditors to be conservative and to be able to clearly justify their decisions given the greater societal and economic consequences of audit opinions. In this sense, the theory-driven approach could be a better fit for auditors to apply big data in auditing (Alles, 2016). For one reason, auditors gain more confidence specifying to at least some extent what they are looking for. Moreover, a conceptual framework within which auditors can develop expectations can improve and simplify the identification of evidence patterns in big data and contributes to audit

planning and effectiveness. The manager interviewed (interviewee 2) supports this point by saying:

> *"In the initiation phase, we must continue to understand where things come from (...). With this approach, we have to realize where things come from so that we can create an expectation (...) and understand the correlations of the various accounts. Only perceiving the business well and understanding how things are related is that you can formulate a minimally valid expectation."*

In an ABE context, this mindset is also important because when auditors test all of the transactions it could result in an unreasonable number of exceptions (Vasarhelyi, 2015). Thus, when the correlation is not designed correctly, it could result in thousands of exceptions which according to the standards have to investigated.

When considering such a big change on the audit field, the impact of legislation on the profession's analytical procedure choices should not be ignored. Due to their responsibility towards their clients and remaining stakeholders, auditors' practitioners are always under high scrutiny by regulators. In this sense, one of the challenges that auditors are facing is the lack of knowledge of regulators on this topic, partly because international standards do not clearly specify what to do when big data and data analytics are in play. Interviewee 4 mentioned this matter:

> *"Imagine what it is, the regulator itself, the CMCM [7], asking us questions about using the analytics. (...) And they ask questions like, "So, now, does everything come? And how does it all come? Oh, and you do not see files? So, is that reasonable? " What is certain is that there is nothing written about the fact that I analyze the entire population"*

The research tried to disclose the extent to which the actual international standards on auditing (ISAs) support the adoption of the big data and data analytics, as the literature

[7] The CMVM is the entity responsible for public oversight of statutory auditor (ROC), Certified Auditors' Association (SROC), auditors and audit entities registered in Portugal, under the terms of Law No. 148/2015.

suggested, or whether they need to be updated. The opinion on this topic was unanimous. As it was possible to confirm through interviewee 2 statement:

> *"I think there's something in there right now that allows you to do what we do. Otherwise, we would not do it if we were not in ISAs. But leave some doubts in the air. That is, by not being clear, I would say that maybe the ISAs, sooner or later will have to be updated to deal with this matter."*

Moreover, there is some internal push in that direction since the internal audit methodology itself is already adopting a new approach to include big data and data analytics in audit procedures. The senior interviewed (interviewee 3) said the following:

> *"From the EY methodology point of view, I can say that is being created and will come into force, (...) surely in 2019, a EY GAM of analytics. This means that people are giving importance to this subject (...). There are no obstacles in this path, on the contrary, it is almost being implemented by force."*

To preserve a more conservative perspective, the author was challenged to state some restrictions about the extension of big data and data analytics in the auditing procedures. The principle of audit is to give some level of assurance over a company financial statement. It does not imply a 100 per cent guarantee that those financial statements are free of misstatements. That is the reason why auditors have the concept of materiality. So, in fact, auditors have no need to test 100 per cent of all transactions. Another reason is linked to the nature of some transactions, that due its low frequency they do not allow the use of data analytics. Finally, there are some doubts about if auditors will take full advantage of big data by including exogenous information on their data analytics tools. Interviewee 1 said the following:

> *"One of the biggest questions is this: in auditing we use a concept called materiality that allows me to conclude whether something is right or wrong and to make a difference to the information provided by the client. (...) At this moment, I do not believe that it is possible to do a full audit in big date. (...) There is certain information that or I have a set up cost so high or the transactions are of such a special feature that I will never do a big data audit."*

Although, the auditing future does not raise doubts: it will be digital. Mainly because internal politics allied with IT environmental changes that will facilitate the adoption of analytics tools. In this sense, auditors already understand that it is better to be proactive than to be reactive in terms of incorporating big data and data analytics into audits. Nevertheless, this comes up with further questions, such as the organizational structure of a Big 4 company. The partner interviewed (interviewee 4) says the following regarding this subject:

> *"Now the base of the pyramid will have to start to shrink. In the future, in 4 or 5 years there won't be a need to hire so many juniors. Because? Because things are going to be much more automated, what we're going to need is experienced auditors that have task-specific, and client-specific knowledge to conduct better analysis."*

At the same time, auditors can feel the need to gain new skills in order to adapt to new trends. It is one of the points raised by interviewee 5:

> *"As the role of the auditor becomes more strategic, it is only natural that there will be great investment by audit professionals in gaining skills related to data analysis. There is also a growing investment in support teams that present a more technological profile."*

Moreover, EY admits that in the future the adopting of big data and data analytics will imply breaking out of the boundaries that exist around their service lines so the company can go to market with an integrated and sector-focused offering.

Above all, something that the researcher could observe and witness is the full commitment of the organization with this subject. The EY Christmas and New Year's message comprehended the following statement:

> *"As digital continues to impact our working world, we will need to innovate our operating model, which will require a new mindset as part of a greater process of cultural change." (Donato Iacovone, Regional Managing Partner Mediterranean)*

IV. Conclusions and Future Research Directions

Overall, big data and data analytics, whether considered as an evolutionary or revolutionary development in technology, remains a means towards an end and not an end in itself. For auditors to find value in it, big data and related analytics must lead to more efficient, effective and higher-quality audits. Moreover, the implementation of these tools is fundamental to stay competitive in the market for assurance services.

Since this topic is quite new, conducting empirical research can be challenging because of the lack of quantitative data. For this reason, the research approaches have to be based on qualitative methods such as interviews, case studies, surveys, and even focus groups in order to obtain insights and perspectives to answer questions such as: What are the characteristic of big data that have the greatest potential to improve the quality of financial statement audits?; What are the costs of incorporating big data into auditing processes?; What are the technical, business vision, and cognitive skill sets auditors will need in a big data environment?; and what would be the best approach to incorporate specific big data elements and data analytics into specific phases of the audit process?

The main limitation that the author faced during this WP were the limited and scientific literature available about the object of study. Additionally, it is important to mention that the conclusions of this study are limited due to the reduced number of interviews conducted. In part the cause of this last point was due to the low availability of EY employees. A final limitation relates to the fact that the researcher participated actively in the company which may eventually lead to biased interpretations and judgements of the topic.

In addition to the issues raised above, this WP can be used as a basis for future research to increase the number of interviewees and thus reach a more robust conclusion. Furthermore, the author suggests that researches should be done to study the effect of the use

of artificial intelligence on audit processes as well as their role in detecting scans in accounting transactions.

V. References

Alles, M. 2015. “Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession”. Accounting Horizons 29 (2): 439–449. doi:10.2308/acch-51067

Alles, M., A. Kogan, and M. A. Vasarhelyi. 2002. Feasibility and economics of continuous assurance. Auditing: A Journal of Practice & Theory. 21 (1): 125–138. doi: 10.2308/aud.2002.21.1.125

Alles, M., and G. Gray. 2014. “Developing a Framework for the Role of Big Data in Auditing: A Synthesis of the Literature”. Working paper, Rutgers, The State University of New Jersey.

Alles, M., and G. Gray. 2016. “Incorporating Big Data in Audits: Identifying Inhibitors and a Research Agenda to Address Those Inhibitors.” *International Journal of Accounting Information Systems* 22: 44–59. doi: 10.1016/j.accinf.2016.07.004.

Appelbaum, D. 2016. “Securing big data provenance for auditors: The Big Data provenance black box as reliable evidence”. Journal of Emerging Technologies in Accounting 13 (1): 17–36. doi:10.2308/jeta-51473

Appelbaum, D., Kogan A., and Vasarhelyi, M. 2017. “Big Data and Analytics in the Modern Audit Engagement: Research Needs.” *Auditing* 36 (4): 1–27. doi:10.2308/ajpt-51684.

Buhl, Hans Ulrich; Röglinger, Maximilian; Moser, Florian; and Heidemann, Julia (2013) "Big Data - A Fashionable Topic with(out) Sustainable Relevance for Research and Practice?" *Business & Information Systems Engineering*: Vol. 5: Iss. 2, 65-69.

Bynjolfsson, E., J. Hammerbacher, and B. Stevens. 2011. Competing through data: Three experts offer their game plans. McKinsey Quarterly, October. https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/competing-through-data-three-experts-offer-their-game-plans

Cao, M., R. Chychyla, and T. Stewart. 2015. “Big Data analytics in financial statement audits”. Accounting Horizons 29 (2): 423–429. doi:10.2308/acch-51068

Connolly, S., 2012. “7 key drivers for the Big Data market”. Blog Posting May 14th. http://hortonworks.com/blog/7-key-drivers-for-the-big-data-market/

Cukier, K., and V. Mayer-Schoenberger. 2013. “The Rise of Big Data: How It’s Changing the Way We Think about the World”. Foreign Affairs 92 (3): 28-40. Issn: 00157120

Dai, J., P. E. Byrnes, and M.A. Vasarhelyi. 2013. “Audit Analytics: A Field Study of Credit Card After-Sale Service Problem Detection at a Major Bank”. Working paper, Rutgers, The State University of New Jersey.

Deloitte. 2013. “Formula for growth: Innovation, Big Data & Data Analytics”. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consumer-business/us-da-executivesummary_gma_big_data_041913.pdf

EY. 2015. “How big data and analytics are transforming the audit”. https://www.ey.com/en_gl/assurance/how-big-data-and-analytics-are-transforming-the-audit

Glover, S. M., D. F. Prawitt, and M. S. Drake. 2015. “Between a rock and a hard place: A path forward for using substantive analytical procedures in auditing large P&L accounts: Commentary and analysis”. Auditing: A Journal of Practice & Theory 34 (3): 161–179. doi:10.2308/ajpt-50978

Keltanen, M., 2013. Why ‘lean data’ beats Big Data. Guardian Newspaper http://www.theguardian.com/media-network/media-network-blog/2013/apr/16/big-data-lean-strategy-business.

KPMG. 2018. “Big Data Fraud”. https://assets.kpmg/content/dam/kpmg/ca/pdf/2018/03/finding-the-evidentiary-needle-demonstrating-fraud-in-the-big-data-haystack.pdf

Littley, J. 2012. “Leveraging Data Analytics and Continuous Auditing Processes for Improved Audit Planning, Effectiveness, and Efficiency”. http://www.kpmg.com/US/en/IssuesAndInsights/ArticlesPublications/Documents/data-analytics-continuous-auditing.pdf

Mayew, W., and M. Venkatachalam. 2012. “Speech analysis in financial markets”. Foundations and Trends in Accounting 7 (2): 73–130. doi:10.1561/1400000024

Moffitt, K., and M. Vasarhelyi. 2013. “AIS in an age of Big Data. Journal of Information Systems”. Journal of Information Systems. 27 (2): 1–19. doi:10.2308/isys-10372

Moore, G.A., 2002. Crossing the Chasm, Marketing and Selling Disruptive Products to Mainstream Customers. Harper Business, New York.

Public Company Accounting Oversight Board (PCAOB). 2010. PCAOB Release No. 2010-004, Docket Matter No. 026. https://pcaobus.org/Rulemaking/Docket034/Release_2013-005_ARM.pdf

PWC. 2018. “AI will help answer the big question about data”. https://www.pwc.com/us/en/advisory-services/assets/ai-predictions-2018-report.pdf

Rezaee, Z., Dorestani A., and Aliabadi, S. 2018. “Application of Time Series Analyses in Big Data: Practical, Research, and Education Implications.” Journal of Emerging Technologies in Accounting 15 (1): 183–97. doi: 10.2308/jeta-51967.

Richins, G., Stapleton, A., Stratopoulos, T., and Wong, C. 2016. “Big Data Analytics: Opportunity or Threat for the Accounting Profession?” *Journal of Information Systems* 31 (3): 63–79. doi:10.2139/ssrn.2813817.

Rose, Anna M., Jacob M. Rose, Kerri-Ann Sanderson, and Jay C. Thibodeau. 2017. “When Should Audit Firms Introduce Analyses of Big Data Into the Audit Process?” *Journal of Information Systems* 31 (3): 81–99. doi:10.2308/isys-51837.

Ryan, B., Scapens, R. W., & Theobald, M. 2002. Research method and methodology in finance and accounting.

Schneider, Gary P., Jun Dai, Diane J. Janvrin, Kemi Ajayi, and Robyn L. Raschke. 2015. “Infer, Predict, and Assure: Accounting Opportunities in Data Analytics.” *Accounting Horizons* 29 (3): 719–42. Doi:10.2308/acch-51140.

Settles, B. 2009. Active Learning Literature Survey 1648. Computer Sciences Technical Report, University of Wisconsin–Madison.

Tang, Jiali (Jenna); Karim, Khondkar E. 2017. “Analytics: Implications for the audit profession”. CPA Journal. New York State Society.Vol. 87 Iss. 6. 34-39

Vasarhelyi, M. A., A. Kogan, and B. M. Tuttle. 2015. “Big Data in accounting: An overview”. Accounting Horizons 29 (2): 381–396. doi:10.2308/acch-51071

Vasarhelyi, M. A., and F. B. Halper. 1991. The continuous audit of online systems. Auditing: A Journal of Practice & Theory 10 (1): 110–125. Issn: 0278-0380

Yin, R.K. 2014. Case Study Research: Design and Methods, 5th edition. Sage, Thousand Oaks, CA.

Yoon, K., L. Hoogduin, and L. Zhang. 2015. “Big Data as complementary audit evidence”. Accounting Horizons 29 (2): 431–438. doi:10.2308/acch-51076

VI. Appendixes

Appendix 1. Advantages of big data and data analytics implementation and supported research of the Work Project

Benefits	Existing Research
Build credibility	Alles (2015)
Reduce cost	Littley (2012)
Improve processes/better control	Alles (2016)
Reduce risk	Alles (2014)
Increase predictive power	Littley (2012)
Improve fraud detection	Cao et al. (2016)
Predict future tax liabilities	Schneider et al., 2015)

Appendix 2. Interviews' details

Interviewee	Category	Channel	Duration (minutes)	Date
1	Associate Partner	Orally (recorded)	40	04/10/2018
2	Manager	Skype Call (recorded)	40	17/10/2018
3	Senior	Skype Call (recorded)	30	19/10/2018
4	Partner	Skype Call (recorded)	40	22/11/2018
5	Staff (Analytics Team)	Email	30	27/11/2018

Appendix 3. Semi-structure interviews guide

Thesis Interview: Impact of big data and data analytics on the external audit process

1. How do you feel about the evolution of auditing in recent years regarding the adoption of big data and data analytics? In what aspect has the How does the profession of auditor changed? Specifically, how do you think it has changed the way you perform your tasks? In what areas have these tools been applied?

2. What are the biggest (positive or negative) impacts you have seen, on an auditor's profession, with the implementation of these tools?

3. How do you rate the auditors' adoption of big data and data analytics?

4. In your opinion, what have been the main drivers in the development / implementation of these tools?

5. What are the biggest risks and / or obstacles you identify in the adoption / implementation of big data and data analytics?

6. What potential application for these tools do you identify at each stage of the audit process? Or what areas do you think have more potential to be applied? Not forgetting that the big data includes not only financial information but also non-financial information.

7. Do you think, on the one hand, that auditing standards (ISA's) are a support or an impediment to the use of big data and data analytics? Do you think it will be necessary to change them?

8. How do you see big data and data analytics changing the audit profession in the future?

Appendix 4. Recorded Interviews

Interview 1 (Associate Partner)

The topic of my thesis is the impact of big data and data analytics on audit processes. The way I intend to add value to the topic is by interviewing professionals in the field to understand their perspective.

My first question is how you have felt the evolution of auditing in recent years in relation to the adoption of big data and data analytics. In what way do you consider that your profession has changed and how do you adopt these tools in your day to day?

Interviewee 1: In my perspective the big date has affected our profession as auditors in two ways: first it affected our profession as auditors as it allows us to demonstrate or explain greater security in my audit process. Why? So far, the audit is always done on a sample basis. It can be statistical, random, or judgmental, but audits have never been done in a logic of global analysis of all movements and transactions, so one of the first changes is in this perspective. We have the ability to pass from an audit that has the possibility to fail due to sampling, to an audit that at the sampling level does not fail because for a given heading a full analysis is done.

Interviewee 1: To summarize the Big Data Tools you have the following advantages:

Interviewee 1: The sensitivity of proving that the audit is safe. Of course, at the same time, it is not only the explanation we can give our clients, but it also has advantages for the auditor. We know that our analysis is safer because, theoretically, I have analyzed everything. Therefore, this gives me the possibility to make my stakeholders more aware of the audit report because they now know that there is a set of analyses that are not partial and random, that are correct according to the audit standards, but which are integral. This is a first relevance of what is to the audit profession the analysis of large data and the use of tools that allow this systematization.

Interviewee 1: The second moment has to do with the following fact that it is not only very important that we present safety in our work, it is also true that we want to be as efficient as possible. So, if clients have a feature and this feature is a complete database and if I can access that database and do my analysis from it, in theory I'll have to be more efficient. Because I do not have to process his data now to collect and analyze. I can straight away go directly to the analysis because I already know what data is available or I have already had at least a first initial investment in knowing how I can quickly put them in my working papers in a smart way so that I make the best possible analysis of them.

Another point had to see in what way, are we changing the audit or not, because of this big data question and analytical tools?

Interviewee 1: There are some difficulties because all of this is very beautiful when I have my client's databases in a way that I can say are compatible with my analysis. There are difficulties regarding this because in some clients this may not be true. Not having this availability on the part of the client or other situations in which the client for reasons of confidentiality do not want to give me full access to the data, and I am not talking about a client who does not want to give information to the auditor, I am speaking of a client who tells us: "I will not give you access to my databases in a complete way, that is, whenever you need something, you hang onto me and I give, so it's the way I also have control over the information I give you. I do not give everything because otherwise I have no control over anything. " So we have the situation that they can not, the one they do not want and the situation they do not have. That is, not all customers have everything so systematized and available for use. Therefore, the massification of the use of these tools does not yet exist. When I say massification I am saying that it is not possible to make a 100% audit based on the analysis of everything because I often do not have this information available.

Interviewee 1: What is happening is that in areas of systematic and recurring transactions, transactions that do not hide business secrets such as revenue transactions and non-transaction costs of raw materials or finished products, then, in these areas, by their size and in which the customers themselves have them more systematized, to this work of integral analyses. While sporadic transactions, even if we have access to them, we can not apply this because we can not enforce a standard.

Interviewee 1: However, neither the first nor the second is to audit. To audit in theory is to draw conclusions if that information was captured and analyzed correctly. So, what I have to do is test that information. Number one, in many situations I also have a set up cost very relevant to get from the extraction of information to the test of information. Number two, there is certain information that I have a set up cost so high or the transactions are of such a special feature that I will never do a big data audit. For example, reading the minutes, I will never be able to process the minutes here, I'm going to ask a tool to read me the full minutes and call me attention to a set of keywords and I'm just going to analyze these keywords or context where they are included.

Interviewee 1: That is, the analysis of the totality of the data is the future, it is an advantage, but also includes in it a series of investments. Not the investment in the technology for the analysis in processing, but the investment itself on the part of the auditor in the attempt to try to create programs of work that are conducive to obtain some security of audit of that same analysis. Because otherwise I can take the risk of obtaining only a summarization of information exclusively and this is not audit.

Does this raise, in your opinion, a problem about whether or not information is reliable, or at least have a greater risk of being manipulative? This because we are working with documents in digital format or, contrary to the past, where we had paper evidence that was harder to change.

Interviewee 1: The auditor must take the most care in analyzing all the information provided by the client. This for a very obvious reason, if the information on which I base my tests is tampered with, my whole analysis will be tampered with. I have to make sure my starting point is correct. While in the analyses I would say traditional auditing, as I go through the tests, I go

indirectly verifying the quality of the IPE[8]. When we are talking about using big data and we use tools that systematize this question, which is exactly one of its advantages, due to its massification this first test done to the IPEs is not made with human or artificial intelligence, reason why, if I have had some flaws in IPE a priori all reviews will be tampered with. And as the analysis is 100% it has another disadvantage being it is going to be 100% wrong. This is one of the great risks that we have in this analysis of mass information, good or bad, we have to audit the initial information that gets into the black box and then is processed. This is one of the risks that we are continually asked to be careful. If I look at auditing in the past and if I look at auditing in the present, one of the main topics that is being raised in the quality of authorship is that of testing the IPEs.

With regard to risks and obstacles, what would you say are the greatest risks in using these tools and obstacles to implementation?

Interviewee 1: Not being a problem, it is a justification for still being in a relatively embryonic situation of the big data audit. The following fact, theoretically, the big data has is the ability to create variable correlations. But we already hope that these variables are related as we want the data to prove this expectation with a certain standard. Putting this in auditory for dummies, I'm expecting a sale to have as a counterpart a customer account. Nowadays, it is very difficult for a sale to have in return a cash account, there are few cash sales. I have notion of what to expect so if there are too many cash sales it is a change to this standard. If I am analyzing the totality of my sales, I can say that in principle, everything that is sales in return for customers follows a certain pattern and sales in return for cash follow a less common pattern. So, if you have a high sales value against cash I have here a change to the standard that I should analyze. Sales in return for any other heading are clearly exceptions and so I will need to get explanations for these exceptions. There may be, I may have revenue in return for an increase in income.

Interviewee 1: The question is how do I train my machine that systematizes this data to analyze or call the attention of these "false hits"? The auditor himself perceives or has to perceive how to program his analysis to get to audit questions and then conclude.

Interviewee 1: And that requires a large initial set up, for example, it requires a set, not only at the level of workflows or flow diagrams of transactions, but at the same time we have to look at the data in two dimensions. By normal physical or procedural analysis, check who approves, who does not approve, who registers, who does not register, but at the same time, check what are the criteria of the application flow charts themselves. For example, I expect all sales to be in return for customers because the system does so, but then if no invoices are issued, the system itself rates that transaction as revenue accruals. I have to realize this because I could have all sales by customer counterparts and find that the standard was 100% accurate and it was not true. I have to realize the specificities of data processing, which the tool itself and the client's application does.

Interviewee 1: This is a relevant initial setup. On the other hand, I'm giving examples that are easier. Now the analysis of big data is not only the analysis of correlations. It can be the analysis of patterns or projections based on past data. And for this, I have to create tools that are intuitive in processing analysis because, for example, I speak for myself, I know how to analyze the data,

[8] IPE means information provided by the entity.

but I do not know how to program to process the data and this is also one of the big data challenges.

And how did you perceive the audit future with these tools?

Interviewee 1: The future of auditing that we want very much does not seem to me, at least at this point in the capacity to do so. The analysis of many data and the use of tools that allow us to systematize this data and analyze this data in patterns is an advantage.

Interviewee 1: What happens is, as in everything in life, there are many enthusiasts who say that auditing could be done all this way. My vision is different, I think it's a great advantage. I'm enthusiastic about, but I do not think it's possible to audit all this way. In a few years I do not know.

Interviewee 1: An interesting topic, which I do not know how we will integrate, would be to incorporate artificial intelligence into a data analysis system. These are algorithms that can detect transactions that go beyond the default. This is the future.

Interviewee 1: In the present form, clearly, the great challenge is how to efficiently incorporate the treatment of big data and data analytics to ensure the profitability of our work, to be able to give the final customer, to the stakeholder of our audit report, a greater security, a greater seal of assurance, a greater facility for them to realize that the analysis is correct and simultaneously I myself can perceive it. Because one thing is to say, these data, are correct I have tested, I will pick up on these data and process them. And then I have a series of analytical relations here. My question is whether these data and these processing I have made are the right ones or not? This is the audit challenge.

The big data does not include only financial information but all the available information. Would you say we can use this information for speculation? For example, how does a vendor's reputation affect business activity?

Interviewee 1: We as good auditors we are, always think of the big data as having all the customer information and now we will do a series of reviews or something will do the reviews for us and then we can conclude about this. The truth is that the work of auditors often also involves creating expectations about a certain behaviour of a particular rubric, and this behaviour may be endogenous when I say, that is, if it is in accordance with what I know. But it may also be exogenous, that is, if the behaviour of my accounts, I will call it that, agrees with the great macroeconomic indicators that also affect them. For example, if I am analyzing why the value of my sales has increased according to the price effect. And my sales are for aluminum or copper so probably the variation in sales is related to the price change in the international commodity exchanges. Therefore, my own audit analysis has to incorporate expectations based on exogenous sources of information. We already do this, but it is not done in a systematic way. That is, today, the treatment of big data is done in a very tight logic. It´s limited to the processing of customer information.

Interviewee 1: At EY we have a great focus on digital auditing, in this focus we have no tool that makes use of different data sources to obtain audit conclusions. The auditor himself already takes into account exogenous information, for example, if inflation has increased X and if I am charging rent, and if there has been an increase in rents, and these depend on the consumer price index, so they are expected to increase consonant this rate.

The last point I wanted to discuss with you is about auditing standards. How can we use these tools or not? If these are in accordance with the standards and in what form do they help, or otherwise prevent the use of these tools?

Interviewee 1: The audit always starts from the principle that you analyze something and come to the conclusion that this something is correct, or not, based on the risk that you define. With these “ifs” that are not bad, these “ifs” are exactly what makes auditing a process that is not merely replicating the transactions to verify if the calculations are well done or not. It is more a process in which I analyze if those transactions make sense according to accounting standards and the reality of the company, and if they make sense to the own audit tests that I do. Standards have a number of concepts of subjective interpretation that have to be taken into account when doing an audit.

Interviewee 1: The big data analysis removes this concept of subjective interpretation. There are some types of assertions, which are the assertions I'll call as more factual, which are much easier to test for by big date and for data analytics tools. Answering questions such as "Has it existed or not?", "Is it true or is it not true?" makes it easier. Now, has it been or not valued? Or whether it is well-presented or not, in the financial demonstration, this is a little different. Because the audit rules are in the background standards to tell if the financial information was presented according to a certain accounting standard and that it fulfilled a set of presentation rules and for this, I have to test a set of assertions. A lot of them always have interpretations so at this time, at the development level we are, I will not be able to do a full audit using big data tools or data analytics, even because I often only have one such transaction.

Interviewee 1: For assertions like currency and completeness, these tools would have much greater weight since they give us greater certainty and test in a more integral way. For the other assertions I find it very difficult at this point to do this analysis. Auditing standards should not be adapted to the fact that I have these tools. Whoever has read the international auditing standards realizes that they are a set, not of rules, but of principles. And like everything in life, when one speaks of principles, the way for interpretation in opened. However, I agree that it would make sense that another ISA be created that would specifically address the care and issues that should be addressed in the big data analysis.

Interviewee 1: One of the biggest questions is this: in auditing we use a concept called materiality that allows me to conclude whether something is right or wrong and to make a difference to the information provided by the client. With the big data, I now can analyze all the data, and despite some discrepancy, say that everything is right and in fact I have an exception there, but this exception is inferior to the materiality. In that case, can I say that everything is all right? My conclusion will be a strange conclusion. According to the rules of auditing everything is alright. There are no materially relevant aspects. But because I have analyzed all the data, I can not say that everything is materially right, it was in an incongruence. Because if I analyzed everything 100%, I should be able to say with certainty if it is right or if it is wrong. At this moment I do not believe that it is possible to do a full audit in big date and therefore I will not be able to completely change the standards. For this reason, I do not believe that the adaptation of standards can be complete. However, at this time, in the way the audit standards are, I also can not pass on an audit opinion when I tested something at 100%. I can only say that I performed a 100% test and presented my findings to my stakeholders....

Finally, how do you think the auditors' adoption curve for these tools is?

Interviewee 1: EY is investing much more in the use of these type of tools than PWC or KPMG. We are not as advanced as the international bosses would like us to be, but EY is certainly ahead in this development. This was due, in part, to the various difficulties already mentioned such as the lack of information on the part of the clients or even by our mindset. We still felt somehow suspicious in terms of just a big date review and we ended up replicating the work. Because at the end of the day we do not want to make mistakes or to have a job that is not supported. At the degree of implementation, I would say that for what I know in the market, our degree of implementation is far superior to the other auditors. Concern is growing this is a disruptive area for auditing. The full data analysis for many areas of audit is the future, I have no doubt. We auditors are very pressured to be early adopters because it is realized that being well done brings us advantages, but I have no doubts about that. I am more afraid of implementation, but this fear does not undermine the advantages I know these tools bring to us.

Interview 2 (Manager)

My thesis is about the impact of big data and data analytics on external audit processes and the way I intend to add value to the topic is to interview professionals in the field, to understand their perspective on the subject from the auditor's point of view and consequently, to perceive how the customer is affected.

Starting with a more general question, how do you think the audit has evolved regarding the adoption of new technologies, not just big data and data analytics?

Interviewee 2: I think we've had significant changes since the time of Ramiro, Madalena and Cristina, that is, 15/20 years ago: you did audits in which you spent much more time on the client. In the past, you have to fill the Leads[9] on paper, folders in the notebook, basically, you would make a sheet and document everything in writing on paper. Then, as time passed, I remembered the stories, there was a personal computer per team, where one of the members documented everything on it but there was still a lot of paper support. Now it has evolved into what we have today, the opposite of what it was like, not so many years ago. What I would draw from a more general point of view was that I think the evolution of the audit has been very quick compared to other types of businesses and industries that have a slightly more time-consuming evolution.

So, you would say that the audit has followed this evolution, in other words, there is a certain delay?

Interviewee 2: No, I would not say that. I even think we're a little ahead. For example, we have now started talking about RPAs (robot process automation) and block chain. We already have RPA to do some things. Basically, you have a computer, you press a button and he does some things for you by itself. It's like a macro but a bit more advanced. For example, he goes to your email, gets some invoice you have received, fills out an EY expense report for you, has it printed so you just have to go get it and file it. Therefore, we are neither ahead nor behind, we are following what is happening in the rest of the world and in the rest of the industries. This is because nowadays we have a much more automated process than what we had 10 years ago.

[9] Leads are the tool EY auditors use to disaggregate balance sheet and P&L rubrics.

How do you think these tools influenced your daily life and the tasks you do? And in the same sense, what areas do you think these tools are the most useful?

Interviewee 2: They are more useful in several ways. The main ones are the fact that you can analyze a big date chain, that is, you can analyze a much larger amount of information than you could have done a few years ago. And on the other hand, by analyzing the totality of the information you can do what we do with the analytics: creating relationships between accounts, between launches, between users and gain some additional comfort by the fact that although you are not seeing the whole population in detail you can have the comfort of making sure everything is connected. Then at the beginning you have to test the old good way by seeing the documents. Only in the middle of the process can you have any additional security because you have analyzed the population and not a sample. It allows you to handle much more information and in a much faster way. It allows us to have a more global and detailed vision. From the audit point of view, it brings us advantages because it increases the quality of our work. On the one hand, it allows us to be more efficient and allows us to dedicate that time, which used to be dedicated to treating files and waiting for excel to work, to make certain analyses to produce information with better quality and with more value to our customer. I do not know if we are already there because these processes always take some time but that is what is expected to happen.

So, in your opinion, one of the main impacts that these tools can bring is in relation to the quality and efficiency of the processes?

Interviewee 2: Yes, they are also those we have been discussing in recent years. The increase in the quality of the audit evidence in the process and the fact that we can better the efficiency of the teams.

Now you can analyze populations at 100%, as you said before, this also implies the change of the mindset of causality where you understood all inputs behind your research for now you have correlations, you only have an output that is an account that relates to another.

Interviewee 2: I do not think so. It will impact the quality, but not in that sense because the "causes" you will always have to see before uploading data wherever it is. You have to understand what causes the processes. Therefore, we divide the SCOTs (Significant classes of transactions), which is what originates the accounting entries in the significant accounts, in several phases, one of them is the initiation, then the process itself, the execution, and then the conclusion. But in the initiation, we must continue to understand where things come from, and with this we can realize what the causes are and where they come from. With this approach, we have to realize where things come from so that we can create an expectation about the result we expect in the future and understand the correlations of the various accounts. Only perceiving the business well and understanding how things are related is that you can formulate a minimally valid expectation. This means that initiation will always have be done and you will always have to recur to more traditional procedures such as substantive procedures for documentary analysis. This also applies to the final stage of the process. For example, if you are talking about sales, it is at this stage that you will ensure that all of it closes. So, you'll see 25 receipts, for example. While at the beginning you will see 25 orders. The invoices are already in the middle in the correlations, you will not see them. What big data and data analytics allow you to do is simplify this process because we gain additional security in another way. So do not stop doing a part of the work in a more traditional way because at the beginning all transactions

are documented and it is very important that we validate this. But what you do is start auditing in a much smarter way and it gives you, in theory, much more security. We must start with the initiation and the cause in order to be able to conclude on the consequences. For example, if you think about journal entries and that relationship that exists between accounts is something I ,5 years ago, had to do in excel with filters, copy to sheets next to me, as if we had to get the GL reports, then people did not even know where it came from, and it was a very manual job, where they lost a lot of time. I now just write what I want in a pop-up and that gives me a result. It means that the process has been simplified and more and more we use tools that analyze the whole information. Another thing is that we also have that Leads program to analyze the fixed assets. You do not waste too much time to put a balance sheet and a property record there and that gives you back the work done. It calculates the amortizations, compares the rates with the previous year, incorporates the 25th decree of 2009. A series of work that you would take a morning or a day to do. Especially if you have several companies there. For example, in Andorra, you have 10 companies, we can load the 10 registrations through the society next door and it does the work all at once.

Entering the field of potentialities, what I have observed in the company, is that we always resort to financial information to carry out this analysis and this information is always made available by the client. But the definition of big data includes not only financial information, but also non-financial information and that comes from various types of sources. In this sense, one of the most simplistic applications would be to use these tools to calculate expectations. For example, an expectation of electricity spending could use external information to collect the price of electricity and through company history to calculate an expectation for December in an October review. Not only in the field of expectations, what are the potentialities you see in implementing the big data?

Interviewee 2: If we go to non-financial, of course they are important data for us from an audit point of view. But we can separate what is useful data in the audit. Although when you receive a diary, what you have there is financial information, we say, using the term that is in the methodology, that we have data points, and you have financial data points and non-financial data points Example of non-financial data points that we use in our processes: the user, the document date, and the release date. On the other hand, value and currency are all financial data points. In my point of view, when I am testing a balances antiquity map or when I am testing the open items of clients, I have to go ensure that that non-financial, which is the document date is in accordance with the conditions contracted with that customer and with the invoice that I issued. In theory, that would make perfect sense not only for electricity but also for diesel and to relate this to the cost of transport. There are a number of options that we could think of in that direction. I think that this information is not readily available, at least for now, and I think it would be difficult for us to validate the reliability of this information. I have to have some assurance that what I am using is in fact what I owe.

What would you say are the greatest risks in using these tools and obstacles in the direction of implementation?

Interviewee 2: A major obstacle to change is people. People are very averse to change. They want to continue doing what they have always done. This is always the biggest obstacle, it is the mentality. It is changing the way you do things. People want to do the same as last year because it is the easiest way to do it at that moment and you do not have much time to think

about innovating. A great difficulty, and from a more practical point of view, the methodology that we have and follow, probably all the Big 4. We now do a digital audit, we have Helix[10] and other work programs, but in methodology, we do not have a digital methodology. And this in a Big 4 that has to deal only in Europe with 28 countries with different regulators that later have different requirements in terms of reporting and regulation. We have to deal with a change in this dimension, in terms of methodology is a major obstacle. We are on the way to have a digital GAM during the middle of next year. We are going to have our normal global audit methodology, which is the one we have today and then we will have a digital, to audit in another way. This is a very big step, in the sense of confronting regulators. Who can say: "you cannot analyze only date without seeing great documentation". For example, our methodology says you do not have to go to key items if you have persuasive or corroborative security, but this is a hard thing to fit. A person who is in the CMVM or another regulator in another country that is already older may also be averse to this type of change and way of working. For example, ISAs also did not have a very large update in the digital sense. International standards governing auditing speak of substantive analytical procedures, but it is a very brief thing.

It was one of the questions I was going to ask you, if you think that, on the one hand, the international standards support or if they are a barrier to the use of these tools?

Interviewee 2: I think there's something in there right now that allows you to do what we do. Otherwise, we would not do it if we were not in ISAs. But leave some doubts in the air. That is, by not being clear, I would say that maybe the ISAs, sooner or later will have to be updated to deal with this matter. But for now, they have not been. It's not something that changes every day. But undoubtedly, I think that as we have a profession that is regulated and that is constantly on subject to great scrutiny, be it by the regulators or the other stakeholders, even the general public when some problems arise. I think there may be obstacles, not only in people's mindsets but even in the technical sense of after all this, if it is supported or allowed in ISAs.

Interviewee 2: And on the client’s side, do you think the client has some role in adopting big data and data analytics, or can it benefit from the introduction of these tools in our work?

At this point, I think the client still does not have much visibility or knowledge on what we can do. Therefore, I have not yet been able to evaluate that very well. Surely when they pay for a service, they will be waiting to have the highest quality service possible. And no doubt that such an audit is better than being 3 weeks “licking” paper and then in the end have managed to see very little, but nevertheless having to finish. Contrary to what we do now. Therefore, quality adds without a doubt. Now, if the client has such visibility, I'd say no, at least at this stage. But in the future, then, it is also possible that clients want to negotiate fees in order to lower the cost of the service. After all, in order to be efficient and to maintain our profession and practice, we have to add something, and I cannot do any consulting or suggestions because I am an auditor. I must be independent.

How do you evaluate the adoption curve of these tools and what would you say are the main drivers?

Interviewee 2: Honestly, I think the main drivers are leadership that is also committed to this subject. Because it is very easy to find resistance and it is very easy for things not to change and everything unfolds as it always has. To change something, I think that this main driver is

[10] Helix Analytics is the previous version of the analytic tool used by auditors before GL happens.

also the commitment of those who are above, of leadership, where tendentially the change is even more difficult. Because they push us in this direction. And this comes of what is produced out there. We also benefit from belonging to a global network. And of investing a lot, a lot, a lot, a lot of money in this type of tools. On the other hand, it is also important, we all in the evaluation have a part in our rate in the last 3 or 4 years connected to what is our performance in the digital viewpoint. It means that if I use the tools in 60 or 70% of my clients, I'm fine. That is, this can be shown, it´s factual. And this link that exists, is the way of leadership to pass the burden of implementing some tools for the ranks of the organization. But of course, even from the auditors' point of view, they prefer to do something clever than to be "licking" paper. It is more stimulating to be doing something that may take a little longer to implement, but which is more challenging from an intellectual point of view.

Interviewee 2: There are a lot of things that are not so good either. To assemble the tools involves drawing the procedures and involving someone more experienced on the team, which not always happens. It means that in addition to being less efficient in the first year, it also implies the greater involvement of people with more experience with the ability to make decisions and complete the work successfully. And the reality is that there is also the lack of availability by everyone to be able to do this kind of work.

Interview 3 (Senior)

The subject for my thesis is the impact of big data and analytics in audit process. And my method to increase value to the subject is to interview some professionals of this area in order to understand their perspective.

First, do you think there has been an effort to implement these tools? How do you apply them into your work and in what context?

Interviewee 3: Indeed, there is a lot of pressure to implement these tools. Since I got in, about three and a half years ago, we didn't use almost anything, they were being implemented. Now it's mandatory, in the part of engagements and, in reality, it is going to become mandatory in almost all of them. The people who work here, either they want it or not, they are going to end up using it. The pressure to use it is a lot and, truth is, it is going to be used everywhere.

But, in the sense of understanding in which areas is more used, how would you say is it used in audit, since its planning to the emission of opinion?

Interviewee 3: In terms of tools, we use GL[11], which is utilized mostly in two stages, the final stage, to help in the substantive tests and in internal control. During the internal control phase, we can resort in GL in matters of fraud and task segregation. Because GL can put together all the information you have in JE's, and that is always an advantage. With GL you can process all the information and not just part of it. And you can also have a wide view of all the areas. Also, if you think about applying GL in internal control, you can also apply it to several affairs, like task segregation. For example, at the same time, you can see if somebody is making spending movements and banks. This could mean that one person is transferring money for yourself. If someone has the power to make movements in their bank accounts and, at the same time the spending account, it can pass the expense between the rain drops. In the end, we use it more in

[11] GL formally General Ledger Analyzers is a data analytics tool used by EY auditors.

the correlations part, which is, in my opinion, the most important part. Basically, you can correlate the accounts as revenue, costumers and cashier and make the cash cycle. It may not tell if the calculation is correct, but it will give you an extra security in the way that you know which percentage of sales generates income. At the end of the day, that's all that matters. If the bank is correct and sales also are, as bad as the company is, in the end not everything is lost.

In terms of potentiality and not forgetting that big data includes financial information and non-financial information, which areas do you think you can explore while using these tools? For example, collect provider's information from a certain company and detect the impact of negative feedback about one of them inside the company.

Interviewee 3: When you talk about big data in financial terms, that is, in audit context is to manage all the JE's. Collecting information outside the company is almost impossible. Evidently, it would useful. That is what we try to analyze when we make our circularization. If you could, automatically have a data base if the provider x, y or z is in insolvency or has problems with payments. Or if costumer x, y or z exist or not exist. That would be useful. Knowing from the start if a provider or a costumer exist would be useful because we could detect fraud cases much earlier. But that would mean that you also need to have gigantic data bases and that you can ensure the reliability of them.

That is correct, I was insisting on this point because, truth is, all the information we use is on the costumer's data base.

Interviewee 3: The data base belongs to them, but you access it and test it. The principle is: whatever you may collect from their data base, you need to check if it is real and if it is right. It's the matter of testing IPE. Whatever they give you, you need to check if it is right.

About risks and obstacles, for example, every month we receive updates from certain software or tool, do you think it's easy to keep up with them or do you feel there is a barrier to their implantation?

Interviewee 3: The main obstacle there is in implement tools is the people. There is resistance to change. Mainly when you are with the costumer and you need to work with older people, no one wants to use the x, y or z software. Nobody wants to use the costumer portal tool, nobody wants to use GL, nobody wants to use anything. Sometimes because they don't have time, sometimes because they can't find their utility. In terms of utility, I think it is incredible useful, it may have more or less time, more or less resistance.

About the risks, you talked about the IPE's tests. Do you think that the quality of information may risk the implementation of these tools?

Interviewee 3: The subject of IPE's is relatively new, regulators have been insisting on the reliability of these data bases. Therefrom, a lot of people have been paying too much attention to the quality of that information. How do you guarantee that the information that is provided is correct and may be used in your analysis? Because your analysis always depends on the information you insert on your tool and all that information needs to be tested first. In a data base, how can you be sure? In the last instance, a client may give you a "hammered" (manipulated) data base and, because of that, all your analysis is wrong. We have been receiving formation for that situation, for example, when you're analyzing the costumer's ageing, you need to check the invoice maturity. Other example is when you input the statements on the GL

software, you need to test the data points, that is, test the dates when it entered the system, among other things.

About international standards, do you think in one hand they support or on the other hand make barriers to the use of big data and data analytics?

Interviewee 3: I'm not familiar with ISA's, but I can tell with certain that they don't impose any obstacle to the use of these tools. From the EY methodology point of view, I can say that is being created and will come into force, supposedly, in the beginning of 2019, surely in 2019, a EY GAM[12] of analytics. This means that people are giving importance to this subject and are going to launch a GAM just for analytics. There are no obstacles in this path, on the contrary, it is almost being implemented by force.

Last question, what is, in your opinion, the biggest impact caused by big data and data analytics?

Interviewee 3: The biggest advantage, I think I mentioned it at the beginning, is that in audit you work with samples because you didn't have, and partly you still don't have, the possibility of treating the information as a whole. You focus on a small parcel and say that you've checked that, so the rest is also right. Sometimes you try to cover as much as you can, but at times it's not possible to much. Imagine you have 100.000 sales invoices, all of them with 100€ in value. Your sales are elevated but you can't see all of those invoices and then you have to work with sampling. With this subject you can treat all your information. And the correlations are that, you can put all that information from sales and correlate that information with cash. This means that I'm able to pick it all up, correlate everything and then have a wide view of everything. This gives more quality to my work, but also more security. Imagine you have 100 million and of those 100 million you only know the origin of some and where did they go. Basically, before you had security over a percentage, now you can see everything. This means that if I say that I put 5 million in sales, I know where they went, however, before I had 5 million, but I only saw 4 million. Now I know that 4 million went to this, 20k went to that, another 5k went to another place and I can see everything. Sure, those 20k or 5k do not matter that much to me, but I can track where they go and how they went. Before you had to look it up first, but now you don't need to do it, just input the information and the system will tell you where the amounts are going, and you just need to check the counterparts. You don't need to go from JE to JE or looking for invoice. And that's the advantage, you can have a wider look into all of the rubrics.

Interview 4 (Partner)

The topic of my thesis is the impact of big data and data analytics on audit processes. And the way I intend to add value to the topic is interviewing professionals in the field to realize their perspective.

My first question is how you have felt the evolution of auditing in recent years in relation to the adoption of big data and data analytics. In what way do you consider that your profession has changed and how do you adopt these tools in your day to day?

[12] EY GAM formally EY Global Audit Methodology is the internal audit standard used in EY assurance department.

Interviewee 4: The introduction of analytics in the audit has been and is still overwhelming. If you do not introduce analytics and robotization tools and everything you can think of so that our auditing process is more automated, and you can analyze all transactions globally, whoever does not do so is lagging behind. And I really felt it, I started in EY and in this profession 20 years ago, and when I went in, no one talked about it. That is, when I entered, I had taken a small excel course and it was the icing on the cake. People were not used to it. There was not a computer for each, it was a shared computer, only the senior one was entitled to one. We worked on paper and everything was done by samples. It is not to say that the audit was not well done, it just did not have the broad vision it has now with the new tools. We are now able to show management, things that in the past were impossible to show. It was not that with the samples of the 25, we did not get that statistical and echometric demonstration. And why the 25, by the statistical demonstration of being more likely to fall into those 25. And we managed to show the administration things that many of them did not see. And they would ask, "Only after 25 how could you detect this?" Because statistically we were able to present results. But from the moment we stop working with statistics and work as a whole, our audit becomes much more reliable. We can draw much more comprehensive conclusions. In such a way that we show to the clients' administration that they do not even know they exist. And we have to walk in that direction. That is, change our own way of looking at audit work. And make it much more current and sexier.

In that sense, what are the biggest impacts, positive or negative, that you have felt?

Interviewee 4: In terms of positive impacts what I think that we have effectively been able to visualize a lot more and perceive the business much better than we realized. That is, being able to visualize how a given transaction starts, processes and ends allows us to have a much more comprehensive view not only at the audit level, but at the level of producing reports for stakeholders. This is undoubtedly one of the great advantages.

Interviewee 4: One of the disadvantages has to do with the team, the audit team is much more seated around a computer and loses the human factor a little, which is to go talk and question people. Because auditing is not just papers. Are we not going to turn into robots? Not really. Auditing is not just about robotization. Robotization will make our job easier but it will not be a conclusion.

As a positive impact, you pointed out that we can visualize the transactions as a whole, how do you think this factor manifests itself in the output of an audit process?

Interviewee 4: It is in terms of opinion safety and quality. By way of example, we gave a presentation to the VMY administration and made a section in the management letter of analytics itself in which we demonstrated to the customer that he had many users touching certain accounts. The administration itself did not realize why it was that it had 60 of its 200 users tinkering with a sales account and tinkering with the inventory account. He could not figure out why in the core business of the company there were so many SAP users who had transacted the account. It is adding value to the customer. The administration itself admitted that no one had ever presented it to them. And it was only through half a dozen analytics queries, nothing special. But of course, the team knew how to walk in search of the key factors. The work itself increased in quality and it was undoubtedly much more secure. We now realize that we have to go looking for other things, not only whether or not there are errors in the accounts

is whether someone made a particular transaction. The ability to detect fraud is far superior at this time.

Moving on to the next question, how do you evaluate the adoption curve on the part of the auditors?

Interviewee 4: Here we have to separate Big 4 from the not Big 4. Not Big 4, forget it, they are light years away, do not use any of it. Zero zero. I know because I belong to the order, I go to the meetings of the order and I belong to the organs of order and I am now going to be part of a study group to talk about this topic because they still have nothing. As for the Big 4, it's not about us, but we are sincerely ahead of everyone. And this is evident. We won *Yellow Submarine* in France in a join audit between EY and PWC, where we won one or won another. And we won because we were able to demonstrate to the customer this kind of analysis, transactions as a whole. That is, for example, that all sales were deposited or that instead a part of sales were not deposited and they did not know that.

That meaning, the use of these tools is already highly valued by the customer?

Interviewee 4: It is highly valued by the customer. Our 2018 audit at *Yellow Submarine* is completely different because we sell it, so they want to show you those reviews every year. The customer himself stated that the others did not do this and PWC confirmed. So now we are ahead. We can´t talk about KPMG and Delloite. So, the only ones that could be compared to us are PWC, the one that has the most similar tools. But still, it does not have such global tools. In fact, we have invested many billions in this kind of tools, in this kind of way of auditing to stand out. Here, we also gain from being global. This is because themes like Canvas, Analytics, among others are developed out there and shared with everyone. This helps us have much more powerful tools. It would be impossible for us in Portugal to do all this alone.

And what do you think are the main drivers for the implementation and development of big data and data analytics in auditing?

Interviewee 4: It is the client, but as a background. First, I think it was due to the position that EY adopted from wanting to stand out and from wanting to be the first. When we did the Vision 2020, the EY wanted, wants and will always want to be the number one in the world and therefore had to stand out. Therefore, they began to make analyses of the area where they could stand out and found that this was a very effective way for customers to see value. Our mantras have always bet on quality. So, this approach was in that sense, the will to be the best. The driver, was mainly internal and in the direction of increasing quality. And in addition, to follow the technological trends that are already applied in other industries. We cannot be left behind. On the other hand, the invoice we present to customers is not getting bigger so we have to find ways to save work. At this point, it is still not a reduction factor because the more we see that we achieve the more we want. That is, when you see that you have a huge potential to know more about the customer, you will not want to save hours, on the contrary, you will want to spend more hours to know more. But there will be a limit, from which you will not want to do more analysis and cut costs. In the future, this will effectively reduce people.

Do you think that on the one hand, there will be a reduction in the number of people or a change of functions of the same? In the sense that they stop doing the work done by the tools.

Interviewee 4: Yes, that will also have to happen. Before, the organizational pyramid was very broad in base because you needed a lot of people to analyze documents and a host of other things that now will not be necessary. "Now the base of the pyramid will have to start to shrink. In the future, in 4 or 5 years there will not be a need to hire so many juniors. Because? Because things are going to be much more automated, what we're going to need is experienced auditors that have task-specific, and client-specific knowledge to conduct better analysis. With this comes an important question, which is: How do we recruit people up if they do not get down there? But that will be another obstacle."

Now, I ask you about the potentialities of big data and data analytics. When I ask you this, I want you to relate to areas, such as the fraud you mentioned earlier. And not forgetting that the big date includes not only financial but not financial information.

Interviewee 4: The use of big data and data analytics will be used across the board, throughout the audit. Therefore, it will be increasingly used in internal control, even in the tests themselves to the control through the analysis of the transactions that are system and in the substantive tests. It will not only be used in the production of the report because it will be necessary to write or issue the certification. It is used in the detection of fraud, this is obvious. We have already analyzed the journal entries, for 4 years it has been mandatory to do this. Even customers themselves are increasingly using these tools to produce management reports, to draw conclusions, for internal investment reviews.

In perspective, what do you think the risks and obstacles to implementation have been?

Interviewee 4: One of the main obstacles is the capacity of the machines, but I think it is combative. I'm not just talking about computers, but also tools. For example, at *Red Shoes*, we use Helix and GL, but these tools cannot incorporate all existing transactions. We have to use the old tools that still use programming codes. So, this is a major deterrent. Moreover, let us not forget that we are investing a lot in this, and therefore we must also include these large companies with a large number of transactions. And that investment is already being made. For example, SAP's ability to generate transactions is huge. Another hindrance in our profession is to be easier to do as before. I think, that there is still, in younger people, aversion to change. It's easier to pick up on what's done and structured than waste time doing basic. But they forget that the quality of the conclusions we draw and the quality that we can deliver to the customer is much better and pays off in the future.

And what is your opinion about the use of information provided by the customer, there is now this issue of testing the IPEs.

Interviewee 4: It's a risk. We must always test the information given to us. But this is an overriding risk. Because there are already alternative processes. Just be present at the time of extraction. While, in the past we had to be the ones to extract the data and it took a week. An example of my experience at *Red Shoes* was that it took a lot of hours and last year it was already possible to ask an outside company working for *Red Shoes* to extract the data. And it took much less time as our computers took hours while in their computers it only took a second. As a result, a great deal of time was spared. It was ridiculous how much time was spent. It would take us a whole week. We would take it out during the night and sometimes you would come in after a week's work and the computer would explode. Just change the mindset a bit and things change. For me this was one of the main obstacles.

Concerning the issue of confidentiality, do you not feel that customers may be reluctant to make all the data available?

Interviewee 4: For some years now, our customers are already used to it. The only thing that exists is that sometimes there are very complex systems of extracting data. For example, in *Yellow Submarine* you have to have a certain application that only exists in France because you can only extract the data there. In fact, the French team had to buy specific servers only for *Yellow Submarine* to be able to extract the data. It takes days on end to extract the documentation. I think that the biggest hurdles are now at a technological level, regarding equipment. Now all our customers have a ERP[13], bad or good. Another client in which we are having difficulty extracting the documentation, but again, it is because of the system and not by the customer's opposition.

The last question I will ask you is regarding the standards, namely the ISAs. Do you think these are an impediment to the use of big data and data analytics? Do you think they need some kind of update?

Interviewee 4: There is no doubt that a rules update is needed. In our company, we are already developing a GAM for analytics. Because so far, our GAM was not ready. The new GAM comes to answer a lot of questions we have. So now the controls are not tested, then what happens to the CRA[14]. Do we get a higher CRA because we do not test the controls? In this sense, the current GAM does not make any sense at all. At most, the CRA should be *moderate*. So here are a number of things that need to be updated. And the ISAs themselves must also juggle. Imagine what it is, the regulator itself, the CMVM, asking us questions about using the analytics. Because what's in the ISAs is that we should use samples. And they ask questions like, "So, now, does everything come? And how does it all come? Oh, and you do not see files? So, is that reasonable? " What is certain is that there is nothing written about the fact that I analyze the entire population. But I do not stop testing checks just because I use analytics. Because for me, I think we have to continue to test the controls. It is the way we can perceive the transactions. Now, testing your controls with analytics is something else.

Interview 5 (Staff)

The subject for my thesis is the impact of big data and analytics in audit process. And my method to increase value to the subject is to interview some professionals of this area in order to understand their perspective. I am only looking for your opinion.

In what aspect has the profession of auditor changed? Specifically, how do you think it has changed the way you perform your tasks?

[13] ERP formally Enterprise Resource Planning is a software is an information system that integrates all the data and processes of an organization into a single system.

[14] CRA means combined risk assessment. Auditor establish one CRA for each relevant assertion for each significant account and disclosure in order to develop an audit strategy that is responsive to the entity's risks of material misstatement.

Interviewee 5: The use of state-of-the-art technological resources in data analytics enables the aggregation of value in the various analyses carried out during the audit process, as it allows the auditor to properly exploit the wealth of information present in the universe of data, rather than small samples.

Interviewee 5: Thus, the analyses are focused on subsets of data selected according to their relevance to the process, improving both the quality of the audit and the value of the insights and conclusions that an auditor is able to provide.

What are the biggest (positive or negative) impacts you have seen, on an auditor's profession, with the implementation of these tools?

Interviewee 5: The adoption of big data and analytics enables auditors to adopt a more strategic position and adapt their approach to the reality of customer data, thereby providing more relevant and accurate information.

How do you rate the auditors' adoption of big data and data analytics?

Interviewee 5: The transformation of audit processes is a reality that gives auditors an opportunity to rethink the way the audit is performed, and like all changes, is not always well received. Many auditors argue that they prefer to continue to do things as they have always done and end up taking a reluctant position on using new methods and techniques, although it being more advanced.

Interviewee 5: There needs to be a realization that even if audits become more automated, "human" auditors will continue to play an important role and that the major focus is to help them and not to replace them.

What are the biggest risks and / or obstacles you identify in the adoption / implementation of big data and data analytics?

Interviewee 5: The problems we often encounter are related to the extraction of the data, which often end up directly affecting the quality of the data. Some problems also arise due to privacy restrictions imposed by the client.

In your opinion, what have been the main drivers in the development / implementation of these tools?

Interviewee 5: The high added value that the use of these tools entails, ends up leading to the natural path being the will to explore more and more the potential of them. Another factor that shows the "willingness" to use these tools is the need to present innovative solutions in the face of competition.

What potential application for these tools do you identify at each stage of the audit process? Or what areas do you think have more potential to be applied? Not forgetting that the big data includes not only financial information but also non-financial information.

Interviewee 5: Any analysis can be leveraged through the application of these tools. I think the great added value here is the ability to gain insights from data from various sources that present

information on various aspects of the business and integrate them so that the analyses present a greater coverage, providing the delivery of more value to the customer.

Interviewee 5: Thus, we have the opportunity to present to the client conclusions that surprise and even fascinate him, instead of simply noticing things that he often already knows.

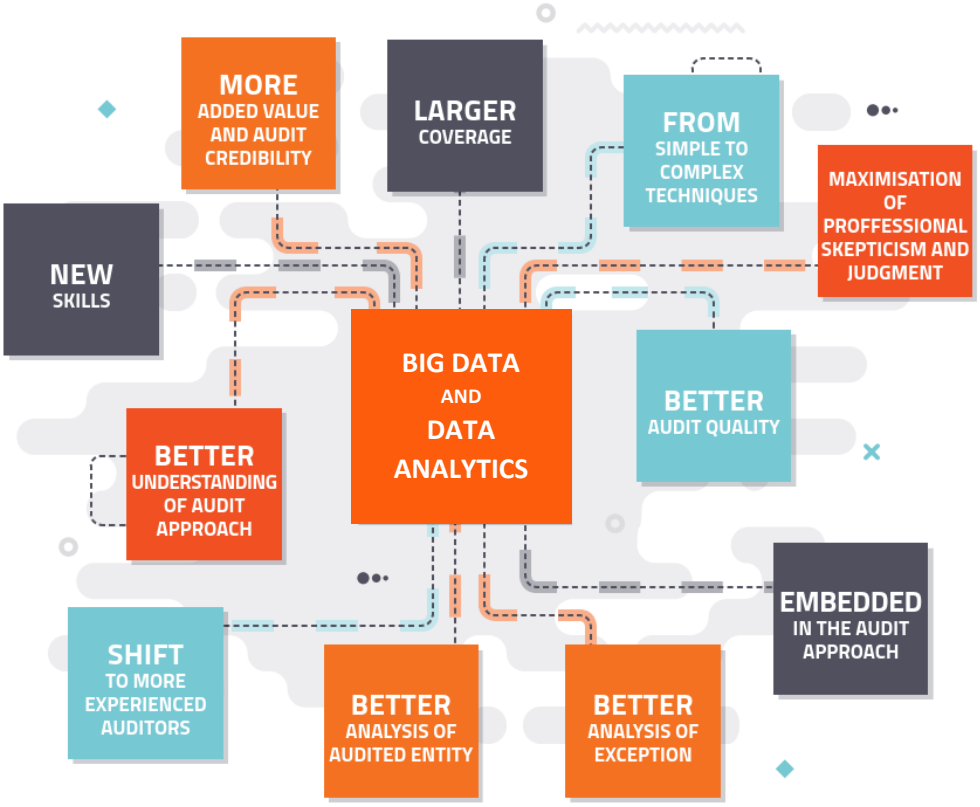
Do you think, on the one hand, that auditing standards (ISA's) are a support or an impediment to the use of big data and data analytics? Do you think it will be necessary to change them?

Interviewee 5: I do not think it's a barrier. I think things can be seen differently. The transformation of audit processes through the use of big data and data analytics has an impact on the way things are done, and they themselves support the use of auditing standards.

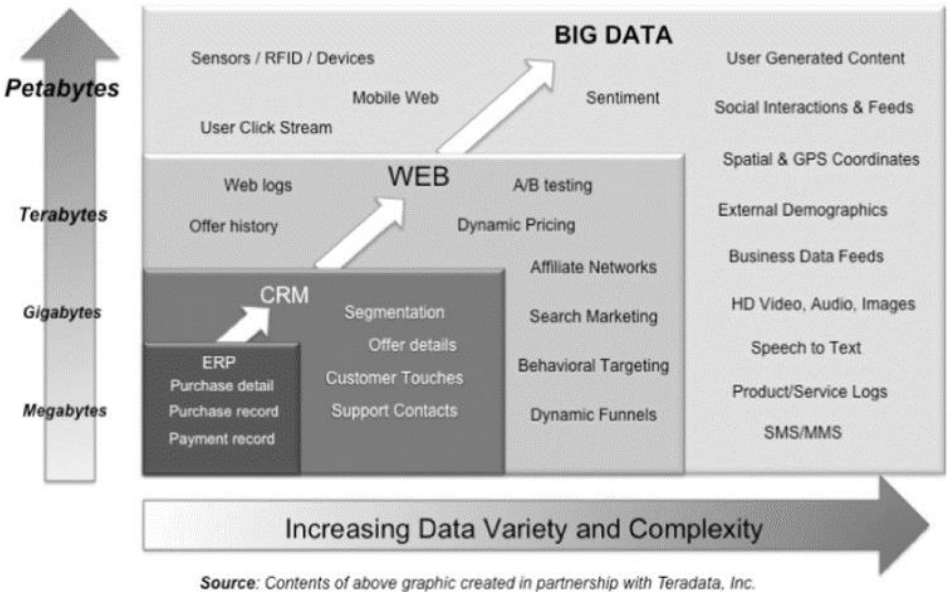
How do you see big data and data analytics changing the audit profession in the future?

Interviewee 5: As the role of the auditor becomes more strategic, it is only natural that there will be great investment by audit professionals in gaining skills related to data analysis. There is also a growing investment in support teams that present a more technological profile.

Appendix 5. Diagram (Adopted from O. Bonhome, E. Gjymshana, M. Jans, D. Kroes, M. Marissen, F. Simpelaere, J. Trumpener and S. Verachtert. 2018. *Data analytics: the future of audit.* Institut des Réviseurs d'Entreprises)



Appendix 6. Connolly's (2012) definition of big data

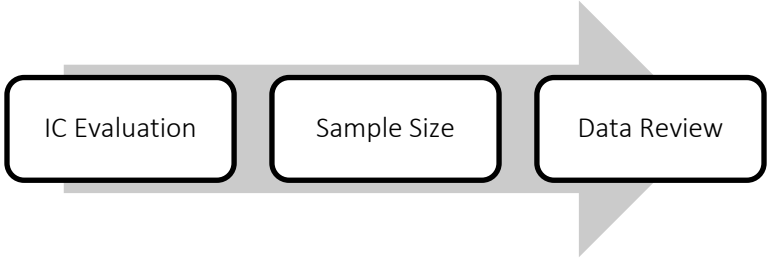


Source: *Contents of above graphic created in partnership with Teradata, Inc.*

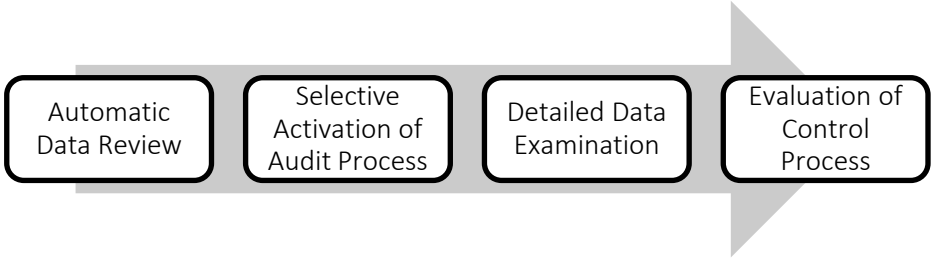
Appendix 7. Alles' (2016) matrix about the paths to expand data analytics in financial audits

		Data Analytic Techniques	
		Traditional (Excel, ACL, Idea)	**Extended** (Visualization, Predictive analytics)
Data Sources	**Traditional** (Accounting & Financial)	A	B
	Extended (Non-Financial Data ➡ Big Data)	C	D

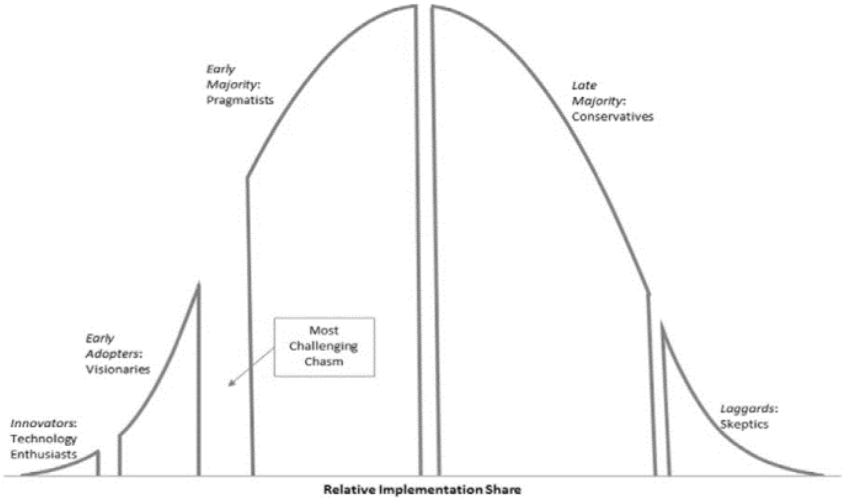
Appendix 8. Current typical audit plan (Adopted from Alles (2016))



Appendix 9. Audit by exception (ABE) (Adopted from Alles (2016))



Appendix 10. Moore's (2002) technology adoption curve



Appendix 11. Richins, Stapleton, Stratopoulos, and Wong's (2017) analytic data flow model